Simulating virtual characters

INTERACTING WITH THE WORLD

New techniques that make virtual characters perform mundane tasks more realistically.

Natural movement of game characters is very important in games and simulations. However, specific tasks such as opening doors or picking up a gun are often poorly animated in games. We have developed techniques that make the character perform these tasks more realistically.

Characters in games generally can do many different actions such as walking, running, or jumping. A common technique to obtain realistic animations is by using motion capture. By tracking the motions of an actor it is possible to have a game character perform the exact same movement as the actor. However, the more actions a character should be able to perform, the more motion capture recordings are needed, leading to huge motion databases. Furthermore, a lot of manual work is required to ensure that transitions between motions look natural. We focus on finding automatic techniques for generating realistic character motion.

"Natural movement of virtual characters is very important in games"

The stepspace

In our technique, motions are generated separately for the upper body (picking up an object) and lower body (walking). The upper body arm motion is created using techniques from robotics. The lower body motion is created by combining and reusing recorded walking motions. After the animations for both the upper and lower body are created, they are stitched back together.

The input for our animation system is a sequence of foot placements, which can be drawn by an animator or calculated automatically from a path that the character should follow. We store a database of walking motions which are automatically separated in different footsteps. New motions are then created by smartly interpolating between these footsteps in the database, combined with a fast parameterization scheme that describes the footsteps using only a few parameters. We then combine the different footsteps to get the final animation. Our technique is fully automatic and needs no editing or post-processing. Furthermore, our technique is real-time. Because we separate the lower body from the upper body, we can now combine upper body motions with the generated motion to perform tasks while walking such as picking up objects or opening a door.

The next step

Many game companies can benefit from this new technique. First of all, the technique can generate animations that walk and pick up objects simultaneously, which is not possible with current techniques. Second, our technique requires less recorded motions. Game companies can therefore spend more time and money on other aspects of the game. We are currently looking into applying a similar technique to the upper body to generate realistic upper body motions based on a small corpus of recorded data. This way, an animator can generate motions based on how a character interacts with the environment fully automatically without having to worry about how the motion is constructed. Furthermore, we are investigating how physical properties can be incorporated in the system, so that the animation is adapted automatically to the constraints of the environment.

Workpackage
3.1 Navigation and Manipulation

Partners
Utrecht University

Budget
500.000 euro

Key Publications

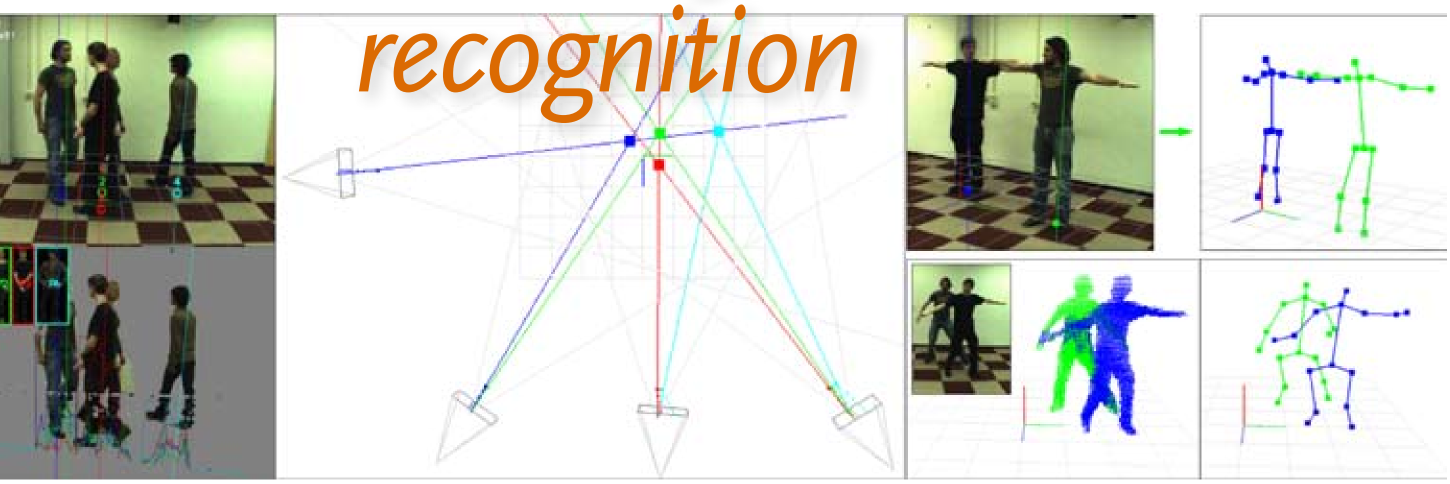
B. J. H. van Basten et al. (2010). Combining Path Planners and Motion Graphs. Computer Animation and Virtual Worlds, to appear.

A. Egges et al. (2010). One Step at a Time: Animating Virtual Characters based on Foot Placement. The Visual Computer, 2010. Special issue of selected papers from Computer Graphics International.

B. J. H. van Basten et al. (2009). Evaluating Distance Metrics for Animation Blending. Fourth International Conference on the Foundation of Digital Games (FDG 2009).

Contact details
Arjan Egges
Utrecht University
egges@cs.uu.nl

Video-based 3D human motion recognition



Interacting with the World

New techniques for 3D tracking of people and understanding poses.

Workpackage
3.2 Detecting, interpreting and affecting user behavior

Partners
Utrecht University
Technical University Delft

Budget
500.000 euro

Key Publication
Luo et al. (2010). Human Pose Estimation for Multiple Persons Based on Volume Reconstruction. 20th International Conference on Pattern Recognition (ICPR), 2010.

Contact details
Remco Veltkamp
Utrecht University
Remco.Veltkamp@cs.uu.nl

A new way to interact with a computer, is to simply use the player's poses, which are tracked and interpreted via cameras. We develop new techniques for accurately tracking people and recognizing their poses, even when they are occluding each other. No special markers on the bodies are needed to track the people.

Although current computer vision research has achieved promising results in interpretating the pose and gestures of a single person given multiple video sequences, intepreting the poses of multiple-people in a relatively dense group is still an open problem. The key difficulties are the inter-person occlusion and limb ambiguities which hamper the interpretation. This project studies and develops new techniques for video-based human pose and gesture recognition. We are developing an efficient and robust platform for multiple-people tracking, body model construction, pose recognition and gesture understanding, all in 3D. We aim at utilizing this platform in human-computer interaction applications like in pose-driven games or gesture-driven presentations.

Beste visibility views

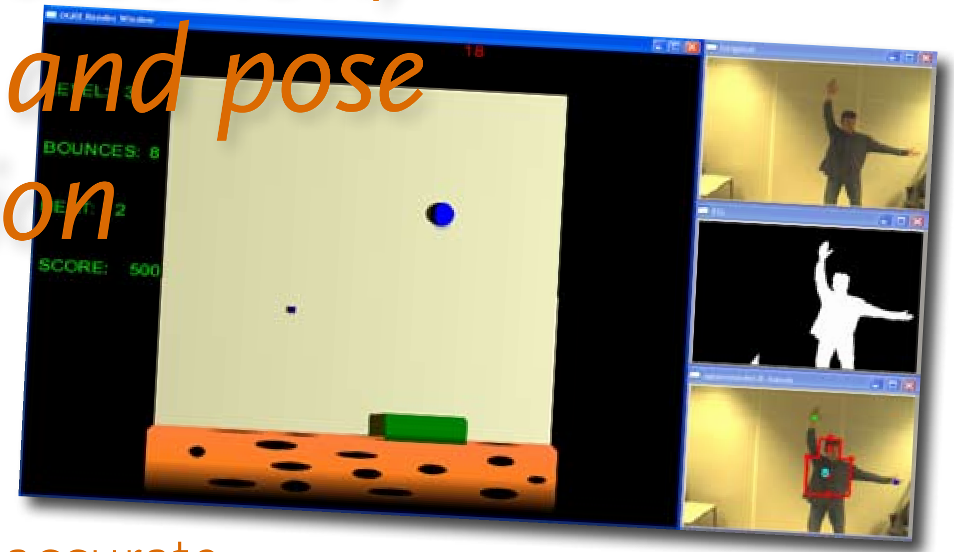
Using multiple cameras, we reconstruct the 3D volume data of moving people in a target scene. The reconstruction is automatic and real time. The volume data consists of the information of the moving people and their poses. The goal is to track the movement and to recognize the poses. To be able to track the movements, first the locations of each individual in the 3D world has to be determined, and then track the location of the next frame by considering the current frame. The basic ideas are (1) to use the appearance of the target person in 2D images, (2) to estimate the 2D location of the person in all views, (3) to backproject the 2D locations from all views onto the 3D world to have an intersection, representing the location of the target person, (4) to employ the location of the previous frames to improve the robustness of the estimation. To overcome the inter-person occlusion problem, a technique based on the best visibility of the views is introduced. The visibility ranking is computed based on run-time measurements of person-person and person-view relative positions. By fusing the information from two views with the best visibility, more robust tracking of people under severe occlusions can be achieved. Having tracked the location of each person, then we can segment the volume data with respect to each person, by fitting a 3D skeleton model to the 3D data. Finally, as a result, the poses of each person can be identified.

"New computer vision techiques will change the way of human-computer interaction"

Gestures and interaction

We will update the tracking methods to a more flexible framework, to enable automatic initialization, and to handle people who enter or leave the scene. We will further improve our prototype multiple-people pose estimation, by integrating more cues. These include the appearance of persons, and motion prediction. We will further extend multi-person pose recognition by using joint locations as feature to classify different poses. While multiple people pose recognition is further elaborated, interpretation of multi-person interaction will be a next step. For the evaluation of our methods, we will develop demonstrators such as a gaming environment with video input, and a gesture-driven slide show presentation framework.

People detection, tracking and pose recognition



Interacting with the World

Fast and accurate modeling human upper body pose.

A well-known video-based application is man-machine interaction, in which people can use their facial expressions, gestures and poses to control e.g. virtual actors or (serious) games. We are developing new methods that allow players to get rid of controllers and play games using intuitive body movements and poses.

Although there have been a significant number of investigations on human motion capture, most of them are marker-based. People need to wear specific suits with markers on it to track the movement of different body parts, which is not convenient for real applications. To solve this problem, marker-less human motion capture system is desired. Compared with a single person situation, multiple person tracking and pose estimation has more challenges, such as dealing with occlusion between persons and self occlusion. The objective of the project is to develop new algorithms wich can detect, track, and model a small group of people in an indoor environment.

Multiple person upper body pose tracking

We propose a real time system which can detect, track people, and recognize poses. In the people detection, tracking and pose recognition system, body parts such as the torso and the hands are segmented from the whole body and tracked over time. The 2D coordinates of these body parts are used as the input of a pose recognition system. By transferring distance and angles between the torso center and the hands into a classifier feature space, simple classifiers, such as the nearest mean classifier, are sufficient for recognizing predefined key poses. The single person detection and tracking is extended to a multiple person scenario. We developed a combined probability estimation approach to detect and track multiple persons for pose estimation at the same time. It can deal with partial and total occlusion between persons by adding torso appearance to the tracker. Moreover, the upper body of each individual is further segmented into head, torso, upper arm and lower arm in a hierarchical way. The joint location and angles are obtained from the pose estimation and can be used for pose recognition.

Mutiple person pose recognition

We will further extend multiple persons pose estimation into pose recognition. The goal is to use joint locations as features to classify different poses. We will also investigate pose detectors to reject non-pose examples based on the proposed features. The approach is to first separate the poses and non-poses, then to clearly distinguish different poses from each other. We will focus on improving the accuracy and robustness of the existing system. The emphasis will on the use of multiple cameras and information fusion. A vision based human pose detection system makes controller free games possible.

"A vision based human pose detection system makes controller free games possible"

Workpackage
3.2 Detecting, interpreting and affecting user behavior

Partners
Delft University of Technology
Utrecht University

Budget
500.000 euro

Key Publications
F. Huo, et al. (2009). Markerless human motion capture and pose recognition. International Workshop on Image Analysis for Multimedia Interactive Services, pp. 13-16.

Contact details
Emile Hendriks
Delft University of Technology
e.a.hendriks@tudelft.nl

Brain controlled navigation through virtual worlds

INTERACTING WITH THE WORLD

New techniques based on brain signals aim to make navigation through virtual worlds more intuitive.

Workpackage
3.3 A Brain Connection Device for Education, Feedback, Gaming, Hands-free Interaction, Joy, Know-how, Learning and More.....

Partners
TNO Human Factors
Utrecht University

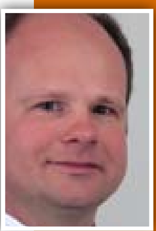
Budget
500.000 euro

Key Publications

Brouwer, A.-M. et al. (2010). A tactile P300 brain-computer interface. Frontiers in Neuroscience, 4:19, 1-12.

Thurlings, M.E. et al. (2010). EEG-Based navigation from a Human Factors perspective. In: D.S. Tan & A. Nijholt (Eds.) Brain-Computer Interfaces, Human-Computer Interaction series, pp. 117-132. London: Spinger-Verlag.

Coffey, E.B.J. et al. (2010). Brain-Machine Interfaces in space: using spontaneous rather than intentionally generated brain signals. Acta Astronautica, 67, 1-11.



Contact details
Jan van Erp
TNO Human Factors
Jan.vanerp@tno.nl

Controlling your movement through a virtual world can be a cognitively demanding task. To make navigation more intuitive, controllers based on hand and body movements such as the Nintendo Wii were recently introduced. The future generation of game controllers aims to be even more intuitive by directly translating brain signals into navigation commands.

Traditional navigation interfaces such as joysticks and gamepads are often not intuitive. This means that users either have to invest many hours of training or allocate significant cognitive resources to the navigation task, reducing overall task performance. To make navigation more intuitive, this project looks into the possibilities of using brain signals for navigation. Our goal is to implement hands-free navigation in at least three dimensions (left, right, rotate). The ultimate goal is that the user's cognitive resources could be fully used for the content of the game instead of the interaction with the interface.

Brain Computer Interfaces

Brain Computer Interfaces (BCI's) enable direct communication between the brain and a computer and come in many different sorts. We are developing passive BCIs: BCIs that use the brain's reaction to specific probe stimuli. The advantage of passive BCIs is that they do not require training but tap into the normal responses of the brain to for instance stimuli that are of particular interest. However, these brain responses are still under voluntary control of the user, making them well suited for BCIs. We explore several types of probe stimuli and brain responses. One of these brain responses is the Steady State Evoked Potential (SSEP). SSEPs are induced by probe stimuli that for instance flicker with a specific frequency. The flicker frequencies can be distilled from the brain signal. When multiple probe stimuli (each with a different frequency) are presented, the user's attention affects which of the stimuli has a stronger effect on the brain signals. In other words: the user can choose from several options by paying attention to one of several probe stimuli. So far, we have developed and tested a BCI based on a visual SSEP. However, the disadvantage of visual probe stimuli is that they require eye movements and may interfere with the visual game environment.

"Future generation game controllers will be hands-free"

Is navigation just a matter of gut feeling?

To overcome the disadvantages of a BCI based on visual probe stimuli, we started to develop BCIs based on touch stimuli. The sense of touch is often underutilized in gaming and does not require (the equivalent of) eye movements. We are now looking into the feasibility of using tactile stimuli with different vibration frequencies as probe stimuli. Of special interest is a set-up in which the probe stimuli are placed inside a belt worn around the user's torso. Choosing left, right, forward, and backward may become very intuitive this way, bringing us a step closer to our goal of hands-free, intuitive navigation. We are also exploring the possibilities of combing visual and touch stimuli to see if this increases the speed or quality of the BCI.